UNIVERSIDADE FEDERAL DO ABC

Programa de Pós-Graduação em Economia

Exame de Defesa de Mestrado | (15/07/2022)

**THIAGO DO CARMO NUNES**

**COMPARAÇÃO DA PERFORMANCE PREDITIVA DE MODELOS ARIMA, VAR E VEC EM SÉRIES TEMPORAIS COINTEGRADAS**

São Bernardo do Campo

**2022**

**THIAGO DO CARMO NUNES**

**COMPARAÇÃO DA PERFORMANCE PREDITIVA DE MODELOS ARIMA, VAR E VEC EM SÉRIES TEMPORAIS COINTEGRADAS**

Exame de Defesa apresentado ao Programa de Pós-Graduação em Economia da Universidade Federal do ABC como requisito final à obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. **Guilherme de Oliveira Lima Cagliari Marques**

São Bernardo do Campo

**2022**

Sistema de Bibliotecas da Universidade Federal do ABC

Elaborada pelo Sistema de Geração de Ficha Catalográfica da UFABC com os dados fornecidos pelo(a) autor(a).

|  |
| --- |
| Carmo Nunes, Thiago  COMPARAÇÃO DA PERFORMANCE PREDITIVA DE MODELOS ARIMA, VAR E VEC EM SÉRIES TEMPORAIS COINTEGRADAS  Thiago do Carmo Nunes – Julho/2022  50 folhas  Orientador: Guilherme de Oliveira Lima Cagliari Marques  Dissertação (Mestrado) – Universidade Federal do ABC, Programa de Pós Graduação em Economia, São Bernardo do Campo, 2022  1.Econometria. 2.Séries Temporais 3. Finanças |

|  |
| --- |
| **Esse exemplar foi revisado e alterado em relação a versão original de acordo com as observações levantadas pela banca examinadora no dia da defesa, sob responsabilidade única do(a) autor(a) e com a anuência do(a) (co)orientador(a).**  **São Bernardo do Campo, 15 de julho de 2022**  **Thiago do Carmo Nunes**  **Nome Completo e Assinatura do(a) autor(a)**  **Guilherme de Oliveira Lima Cagliari Marques**  **Nome Completo e Assinatura do(a) orientador(a)** |

**SIGAA – Sistema Integrado de Gestão de Atividades Acadêmicas**

**UFABC – Fundação Universidade Federal do ABC**

**Programa de Pós Graduação em Economia**

**CNPJ 07.722.779/0001-06**

**Alameda da Universidade, s/nº - Bairro Anchieta - São Bernardo do Campo**

**ppg.economia@ufabc.edu.br**

**FOLHA DE ASSINATURAS**

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Qualificação de Mestrado do candidato, THIAGO DO CARMO NUNES realizada em 15 de julho de 2022.

**AGRADECIMENTOS**

Dedico esse trabalho a minha mãe Aparecida que em todos os momentos me apoiou e me deu forças para iniciar e concluir esse trabalho. Ao meu pai Geraldo que sempre foi meu exemplo de vida. Aos meus amigos de trabalho que desde o início do projeto me deram forças para concluir. Ao meu orientador Guilherme por toda sabedoria transmitida e por estar sempre presente. Ao professor Ramon Garcia que desde o início dessa jornada foi um grande facilitador e coordenador do mestrado em economia da UFABC.

**RESUMO**

O objetivo desta dissertação é comparar a capacidade preditiva dos modelos univariados ARIMA com os modelos multivariados VAR e VEC sob a hipótese de cointegração. A análise foi baseada em simulações de Monte Carlo utilizando métricas estatísticas para avaliar a performance preditiva dos modelos. Para validar as conclusões obtidas via simulações foi realizado um experimento de modelagem com séries temporais reais de cambio (USD/BRL) e inflação acumulada em 12 meses (IPCA).

As simulações mostraram que a capacidade preditiva dos modelos em dados fora da amostra de treinamento é em média equivalente dentro dos intervalos de confiança de 95%. No experimento com dados reais a capacidade preditiva dos modelos foi similar dentro dos intervalos de confiança da média, alinhado a conclusão obtida via simulações.

Os resultados também se mostraram alinhados com as práticas de mercado onde mesmo sabendo a especificação teórica mais adequada ao conjunto de dados no treinamento, deve-se atentar para experimentos de performance fora da amostra para maior assertividade nas predições. Contudo, a amplitude dos intervalos de confiança aponta para a necessidade de realizar validações mais complexas previamente à escolha da estrutura de modelagem para predição.

A conclusão do trabalho foi que mesmo em um conjunto de séries cointegradas, onde pela literatura econométrica existe uma orientação para modelagem via VEC, para maior assertividade das predições é necessário fazer experimentos de validação fora da amostra para escolher o modelo com maior capacidade preditiva.

**ABSTRACT**

This work aims to compare the predictive power of the univariate ARIMA models with the multivariate VAR and VEC models under the cointegration hypothesis. The analysis was condutec on Monte Carlo simulations using statistical methods to evaluate the predictive power. A model was estimated using exchange rate (USD/BRL) and 12-month cumulative inflation (IPCA) to check the Monte Carlo simulation insights.

The model's predictive power on out-of-sample data is similar to Monte Carlo simulations' 95% confidence intervals. The conclusion of the exchange rate and inflation data experiment was in line with simulations. The predictive power of the models was similar within the confidence intervals of the mean. However, the size of the confidence intervals demands more complex validations before choosing the modeling specification for prediction.

The results are also in line with market best practices, where even knowing the appropriate theoretical model specification, we should pay attention to out-of-sample performance experiments for greater assertiveness in predictions.

The work concluded that even in a set of cointegrated series, where there is a guide for modeling VEC in the econometric literature, to get more assertiveness of predictions, it is necessary to do out-of-sample experiments outside the sample to choose the model with greater predictive power.

**LISTA DE FIGURAS**

[Figura 1 - Estratégia de Validação Fora da Amostra Validation 10](#_Toc97798129)

[Figura 2 - Fluxograma do Processo de Simulação 12](#_Toc97798130)

[Figura 3 – MAE | Fixo: n\_obs\_\_100 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90, 95] 13](#_Toc97798131)

[Figura 4 - MAE | Fixo: n\_obs\_\_250 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90%, 95%] 13](#_Toc97798132)

[Figura 5 - MAE | Fixo: n\_obs\_\_500 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90%, 95%] 14](#_Toc97798133)

[Figura 6 - MAE | Fixo: n\_obs\_\_1000 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90%, 95%] 14](#_Toc97798134)

[Figura 7 - MAE | Fixo :n\_obs\_\_100 & n\_ma\_\_0 & n\_fold\_ts\_\_95 | Variando: n\_ts : [2,3] 15](#_Toc97798135)

[Figura 8 - MAE | Fixo :n\_obs\_\_500 & n\_ma\_\_0 & n\_fold\_ts\_\_95 | Variando: n\_ts : [2,3] 15](#_Toc97798136)

[Figura 9 - MAE | Fixo :n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_\_3 | Variando: n\_obs: [100,250,500,1000] 16](#_Toc97798137)

[Figura 10 - MAE | Fixo: n\_obs\_\_1000 & n\_fold\_ts\_\_95 & n\_ts\_\_3 | Variando: n\_ma: [0,12] 17](#_Toc97798138)

[Figura 11 - MSE | Fixo :n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_\_3 | Variando: n\_obs : [100,250,500,1000] 17](#_Toc97798139)

[Figura 12– ECDF | MAE & n\_obs\_\_500 & n\_ma\_\_0 & n\_fold\_ts\_\_85 & n\_ts\_\_2 21](#_Toc97798140)

[Figura 13 - ECDF | MAE & n\_obs\_\_100 & n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_\_3 22](#_Toc97798141)

**LISTA DE TABELAS**

[Tabela 1 - GRID de parâmetros da simulação 8](#_Toc97798142)

[Tabela 2 - Parâmetros fixos da simulação 8](#_Toc97798143)

[Tabela 3 - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_obs\_\_1000 & n\_ma\_\_0 & n\_fold\_ts\_\_85 & n\_ts\_3 | Variando: n\_fold\_ts: [85%,90%,95%] 18](#_Toc97798144)

[Tabela 4 - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_obs\_\_250 & n\_ma\_\_0 & n\_fold\_ts\_\_85 & n\_ts\_3 | Variando: n\_fold\_ts: [85%,90%,95%] 18](#_Toc97798145)

[Tabela 5 - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_3 | Variando: n\_obs: [100,500,250] 19](#_Toc97798146)

[Tabela 6 - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_obs\_\_500 | Variando: n\_ts: [2,3] 20](#_Toc97798147)

[Tabela 7 - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_ts\_\_3 & n\_fold\_ts\_\_95 & n\_obs\_\_500 | Variando: n\_ma: [0,12] 20](#_Toc97798148)

[Tabela 8 – Teste de KS para cada combinação de parâmetros da simulação mantendo fixo o número de observações em 500 unidades. 21](#_Toc97798149)

[Tabela 9 – Teste de KS para cada combinação de parâmetros da simulação mantendo fixo o número de observações em 100 unidades. 22](#_Toc97798150)

**SUMÁRIO**

[1) Introdução 1](#_Toc106614864)

[a) Linha do Tempo 1](#_Toc106614865)

[b) Problema e Objetivo 3](#_Toc106614866)

[2) Revisão Bibliográfica 3](#_Toc106614867)

[3) Metodologia 6](#_Toc106614868)

[a) Design dos Modelos 6](#_Toc106614869)

[b) Metodologia de Simulação 9](#_Toc106614870)

[c) Validação dos Modelos 11](#_Toc106614871)

[4) Resultados 14](#_Toc106614872)

[5) Conclusões 24](#_Toc106614873)

[6) Referências 25](#_Toc106614874)

[7) Apêndice 28](#_Toc106614875)

[8) Repositório 28](#_Toc106614876)

# 

# Introdução

Os modelos estatísticos de previsão são ferramentas usadas por cientistas de diversas áreas para antecipar tendências sobre o comportamento futuro de variáveis como a demanda de atendimentos de emergência (Jones et al., 2002), consumo de combustível de carros em ambientes urbanos (Afana et al., 2018) e para previsões climáticas (Hargreaves et al., 2004). Em Economia, modelos de previsão são empregados tanto por instituições do setor público quanto do setor privado. Entidades como os bancos centrais, por exemplo, utilizam modelos de previsão para antecipar tomadas de decisão acerca da política monetária (Nyoni, 2018). No setor privado a capacidade de predizer um padrão de comportamento futuro traz vantagens competitivas para as empresas como por exemplo aumento da eficiência dos investimentos (Chen et al., 2017).

# Linha do Tempo dos Modelos

Um dos problemas básicos da econometria aplicada às séries temporais é formular modelos de previsão. Box and Jenkins (1970) popularizaram a versão mais conhecida dos modelos de previsão baseada em dados univariados, o Modelo Auto-Regressivo Integrado e de Médias Móveis (ARIMA). Sua proposta é modelar parametricamente a média condicional de uma série temporal baseando-se em seu comportamento passado. Os autores propuseram um algoritmo de fácil implementação contendo basicamente quatro etapas: identificação, estimação, verificação e previsão.

Apesar da análise univariada de séries temporais ser amplamente utilizada na econometria aplicada, as variáveis macroeconômicas tipicamente apresentam comportamentos dinâmicos e interdependentes com outras variáveis. Este comportamento dinâmico é observado em indicadores de atividade econômica, inflação e instrumentos de política monetária aplicados em curvas de retorno de investimentos (Diebold et al., 2006). Para resolver um sistema de variáveis interdependentes (Sims, 1980) desenvolveu os modelos Vetoriais Autorregressivos (VAR). Lütkepohl (1991) explica que a estabilidade dos modelos VAR pressupõe a condição de estacionariedade entre as variáveis integrantes do sistema. A estrutura VAR contribuiu no desenvolvimento e representação matemática de conceitos relacionados às séries temporais macroeconômicas estacionárias.

No entanto, para séries econômicas não estacionárias devido a presença de tendências estocásticas, (Engle & Granger, 1987) mostraram a necessidade do sistema satisfazer uma condição estatística denominada por eles como “cointegração”. Informalmente, séries temporais que apresentam uma relação estrutural entre si compartilhando uma mesma tendência estocástica por meio de uma combinação linear, são ditas cointegradas. Segundo (Enders, 2014) a cointegração possui interpretação econômica como “uma relação de longo prazo entre as variáveis”. (Engle & Granger, 1987) também demonstraram que se um conjunto de variáveis for integrado de primeira ordem e cointegrado, é possível representá-lo sob a forma de um Modelo Vetorial de Correção de Erros (VEC). A principal diferença do modelo VEC em relação ao seu antecessor VAR é a presença de um termo de ajuste de longo prazo, também conhecido como termo de correção de erros. (Johansen, 1988) demonstrou uma forma elegante de testar se séries temporais são cointegradas a partir do modelo VEC.

Um fato estilizado em Finanças é que existem séries temporais onde a causalidade não ocorre de forma bidirecional. Ou seja, apesar dos dados compartilharem uma mesma tendência estocástica, existe exogeneidade em determinado sentido entre as variáveis (Rogalski & Vinso, 1977). Por exemplo, como nos pressupostos oriundos da Teoria Econômica é possível imaginar que uma série temporal referente aos dados de inadimplência de uma pequena carteira de crédito de um banco e a taxa de desemprego do país possam ser correlacionadas. Nesse sistema, pressupõe-se que, apesar de ser possível um compartilhamento de tendência entre as séries, a relação deve ser unidirecional, ou seja, somente a taxa desemprego seria capaz de influenciar a trajetória da inadimplência e não o contrário.

# Problema e Objetivo

Apesar dos protocolos econométricos para estimação de modelos envolvendo dados cointegrados estarem claros na literatura econométrica, em áreas de Finanças (Abdullah & Dwivedy, 2012) e de Riscos do mercado financeiro (Araz et al., 2020) é observado o amplo uso de modelos preditivos para resolver problemas de negócio onde a escolha da sua estrutura foi feita somente comparando a capacidade preditiva de uma lista de modelos em períodos fora da amostra.

A partir dessa observação, este trabalho visa comparar a capacidade preditiva dos modelos clássicos ARIMA, VAR e VEC em conjuntos de dados cointegrados e validar as performances destes modelos em dados fora da amostra de treino.

Para tal, serão analisados os ganhos de performance em relação à capacidade preditiva com séries de dados simuladas. Serão utilizados procedimentos simulatórios de Monte Carlo para gerar os conjuntos de séries temporais cointegradas baseados em coeficientes lineares e termos de perturbação aleatórios.

Para validar as conclusões obtidas via simulações será realizado um experimento de modelagem com séries temporais reais de cambio (USD/BRL) e inflação acumulada em 12 meses (IPCA).

O restante deste trabalho está organizado da seguinte forma: na seção 2 é realizada uma revisão bibliográfica; na seção 3 é apresentada a metodologia abordando o design dos modelos e estratégia de simulação e validação. Na seção 4 são apresentados os resultados das simulações e na seção 5 são descritas as conclusões.

# Revisão Bibliográfica

Em (Fanchon & Wendel, 1992) são comparados modelos VAR e VEC em series de dados cointegrados para modelar preço de gado. A conclusão é que o modelo VAR gerou erros quadráticos médios (MSE) ligeiramente menores do que o VEC. A comparação entre VAR e VEC depende das defasagens incluídas no modelo de previsão, portanto dependendo dos parâmetros escolhidos um modelo performa melhor em relação ao outro.

Em (Bachmeier & Swanson, 2005) são comparados modelos AR, VAR e VEC para modelar a inflação utilizando a técnica de validação fora da amostra para testar a capacidade preditiva dos modelos com dados norte-americanos de 1959 até 1999. Foram também utilizadas simulações de Monte Carlo para gerar estatísticas sobre os modelos. Os autores concluíram que quando os parâmetros do vetor de cointegração utilizado são estimados o modelo VEC sempre tem a capacidade preditiva pior do que um modelo VAR em primeiras diferenças. Isso sugere que trabalhos anteriores que descobriram que os modelos VEC não preveem melhor do que os modelos VAR podem ser em parte devido à presença de erro de estimativa de parâmetro de vetor de cointegração. Portanto isso sustenta a hipótese de que ao encontrar modelos VEC performando pior que VAR em períodos fora da amostra isso pode estar associado ao erro de estimação do vetor de cointegração e não que a estrutura de modelagem como um todo.

Em (Ng et al., 2011) são utilizados modelos VEC para confirmar a influência do capital público no capital privado de construção civil através dos efeitos de gastos públicos sobre o investimento privado (efeito *crowding-in*). O PIB real, investimento do setor público em construção civil e o desemprego são variáveis que influenciam o investimento privado em construção em Hong Kong. O estudo compara regressões lineares (em uma estrutura de defasagens muito próxima ao modelo VAR) com o modelo VEC e mostra através da comparação de performance fora da amostra que o vetor de correção de erros é a estrutura mais adequada para projetar séries reais não estacionárias e variáveis ​​cointegradas.

Em (Anggraeni et al., 2018) foram comparadas as capacidades preditivas de modelos ARIMA e estruturas VAR para prever preços de arroz em conjuntos de dados fora da amostra utilizando indicadores de performance MAPE. No estudo em questão os modelos ARIMA estimados obtiveram uma performance 15% melhor do que o VAR.

(Tehreem et al., 2020) estudam o impacto da temperatura média, demanda de energia, valor agregado setorial e crescimento populacional na qualidade dos recursos hídricos e na taxa de mortalidade no Paquistão. O autor utiliza a estatística do traço proposta por (Johansen, 1988) para mostrar relações de cointegração entre demanda de energia, valor agregado da indústria e a qualidade dos recursos hídricos.

A versatilidade de modelos autorregressivos é comprovada pela ampla variedade de temas aplicados que eles podem oferecer soluções. Foram citados trabalhos no campo da economia, porém na área da saúde tem-se observado um amplo uso dessa classe de modelos.

O ano de 2020 foi de alta demanda para a área de modelos de previsão. A pandemia de COVID-19 exigiu um uso massivo de modelos em áreas de saúde, finanças e setores industriais. Em (Hu et al., 2020) são apresentados métodos de inteligência artificial para prever a duração da pandemia de Covid-19 na China utilizando a validação fora da amostra como método de seleção de modelos. (Salisu et al., 2020) propõe, em sua publicação, modelos clássicos como estruturas VAR sendo utilizados para prever o comportamento dos estoques de petróleo durante a pandemia de COVID-19.

Além da pesquisa de vacinas e medicamentos para mitigar doenças como Covid governos e autoridades centrais utilizaram de modelos autorregressivos para planejarem seus sistemas de saúde para suportar surtos de outras doenças. Para fortalecer as medidas de prevenção e controle da malária do país (Wangdi et al., 2010) apresenta um estudo de modelos de previsão para avaliar a incidência da malária nos distritos endêmicos do Butão usando modelos ARIMA com variáveis exógenas. A recomendação do autor é o aperfeiçoamento e o uso destes modelos no planejamento de medidas de controle do vírus no país.

Além da academia, outros setores da economia utilizam sua infraestrutura de pesquisa e, até mesmo, patrocinam competições de modelos de previsão para entender padrões nos dados relacionados a COVID-19. (Ioannidis et al., 2020) destaca que mesmo com modelos mais modernos para prever dados relacionados a pandemias a tarefa continua sendo árdua visto que existem falhas endógenas e exógenas que precisam de atenção dos cientistas. O trabalho elencou que dentre as razões mais comuns de falha dos modelos de previsão estão: (i) Insuficiência das covariáveis (*features)* e a falta de informações claras sobre o problema; (ii) Premissas erradas no modelo, levantando a necessidade de uso de modelos probabilísticos em apoio as séries de tempo; (iii) Modelos utilizados foram estudados com base em dados de epidemias anteriores, sendo que a disponibilidade e qualidade da informação estavam em patamares muito piores das atuais; (iv) Falta de expertise do negócio (Epidemiologia) em empresas/cientistas que desenvolveram os modelos.

Os aprendizados do artigo são úteis para qualquer área de negócio onde são utilizados modelos preditivos sem uma alta maturidade e amplo conhecimento de negócio onde a seleção dos modelos se baseiam somente em performance fora da amostra de treino.

A partir da revisão bibliográfica, não temos evidências suficientes para pressupor que nenhum dos modelos ARIMA, VAR e VEC possuam capacidade preditiva superior aos demais. O contexto de negócio e as condições de seleção dos parâmetros de cada estrutura são determinantes para escolha do melhor modelo preditivo.

# Metodologia

# Design dos Modelos

Os modelos ARIMA são processos estocásticos constituídos para representar a média condicional de uma série temporal. Esta classe de modelos é utilizada para descrever a trajetória de séries temporais a partir de suas próprias observações e de estimativas de choques em períodos passados. Segundo (Brockwell & Davis, 1991) a estrutura do modelo é composta por três componentes: o componente auto-regressivo, o grau de integração e médias móveis. Uma estrutura ARIMA é representada pela Equação 1:

sendo: p a ordem autorregressiva; d a ordem de diferenciação; q a ordem das médias móveis; Δ operador diferença ; o intercepto; ϕ o coeficiente linear dos autorregressivos; Ө coeficiente linear das médias móveis; é um processo ruído branco.

Os modelos VAR, introduzidos por (Sims, 1980), são estruturas multivariadas representadas por um sistema de equações simultâneas que capturam a existência de relações de interdependência entre variáveis. A estrutura do VAR correlaciona observações atuais de uma variável com suas observações passadas e observações passadas de outras variáveis no sistema. A forma mais simples do VAR é expressa em sua estrutura bivariada de grau um conforme representado pela Equação 3:

sendo: o intercepto; ϕ o coeficiente linear dos autorregressivos; ϵ resíduo ou um processo ruído branco.

Lütkepohl (1991) explica que a estabilidade dos modelos VAR pressupõe a condição de estacionariedade entre as variáveis integrantes do sistema, portanto para um sistema de variáveis não cointegradas e integradas de grau um deve-se modelar o VAR em primeiras diferenças, conforme representado pela Equação 5:

Para os conjuntos de variáveis cointegrados a abordagem econométrica mais adequada é representá-las em um modelo VEC. Segundo (Enders, 2014) incluir o termo de correção de erros no modelo VAR permite equilibrar o ajuste dinâmico de curto prazo e longo prazo entre as variáveis do sistema. Segundo (Engle & Granger, 1987) o termo de correção de erros entre duas variáveis se apresenta na forma da Equação 6 e sua premissa básica é que o termo seja estacionário.

Sendo: os coeficientes lineares do termo de correção de erros; e variáveis aleatórias cointegradas; uma série temporal não integrada, ou seja, estacionária.

Portanto o modelo VEC bivariado de ordem um pode ser descrito a partir da representação da Equação 8:

sendo: o intercepto; ϕ o coeficiente linear dos autorregressivos; o coeficiente linear do termo de longo prazo; ϵ resíduo ou um processo Ruído Branco.

Para este trabalho não foram utilizadas técnicas de otimização de hiperparâmetros nos modelos, portanto todas estruturas de modelagem foram estimados sob a restrição de grau um dos auto-regressivos.

START

Critério de informação bayesiano - BIC O Critério de Informação Bayesiano(BIC) , proposto por Schwarz (1978) é dado por:

onde que é o modelo escolhido, p é o número de parâmetros a serem estimados e n é o número de observações da amostra.

END

# Metodologia de Simulação

Dada a apresentação dos modelos serão apresentados o fluxo de simulação utilizado para comparar modelos ARIMA, VAR e VEC neste trabalho.

Todas as etapas computacionais foram feitas utilizando o software R (R Core Team ,2016). A primeira etapa do processo de simulação é iniciar uma semente para que cada etapa do processo seja reproduzível. Foram definidos conjuntos de parâmetros fixos e uma grade de parâmetros variáveis para a simulação. A grade de parâmetros contém 5 variáveis: (i) **n\_obs**: O número de observações de cada série temporal variando de 100, 250, 500 e 1000 unidades; (ii) **n\_ma:** Suavização ou não das séries temporais via média móvel de 12 períodos; (iii) **n\_fold\_ts:** Proporção da amostra de dados reservado para treino do modelo variando entre 85%, 90% e 95; (iv) **n\_ts:** O número de covariáveis no sistema variando de 2 a até 3 séries temporais.

A grade de parâmetros para as simulações de Monte Carlo está descrita na Tabela 1.

Tabela - GRID de parâmetros da simulação

|  |  |  |
| --- | --- | --- |
| Parâmetros | Rótulo | Valores |
| Número de Observações | ***n\_obs*** | [1000,500,250,100] |
| Suavização por média móvel | ***n\_ma*** | [0,12] |
| Proporção da amostra reservada para treino do modelo | ***n\_fold\_ts*** | [85%,90%,95%] |
| Número de Variáveis | ***n\_ts*** | [2,3] |

Para cada conjunto de parâmetros da grade da Tabela 1 foram gerados 1000 conjuntos de dados para estimação e validação dos modelos. Para os testes de hipótese foi utilizado um nível de confiança de 95%. A grade de parâmetros fixos para as simulações é descrita na Tabela 2.

Tabela - Parâmetros fixos da simulação

|  |  |
| --- | --- |
| Parâmetros | Valor |
| Número de iterações por conjunto de parâmetros | 1000 |
| Semente Inicial do Processo | 20220101 |
| Nível de Confiança usada nos Testes de Hipótese | 95% |

Dado os conjuntos de parâmetros acima, para cada iteração do processo é selecionada uma combinação de variáveis e iniciada a simulação. A primeira etapa da estrutura de repetição atualizar a semente do processo (somada em uma unidade) para mitigar o risco de séries iguais serem geradas. Dada uma nova semente o algoritmo gera o conjunto de dados cointegrado.

O conjunto de equações número 9 simulam três séries temporais cointegradas com um vetor de cointegração. As séries representam as duas tendências estocásticas compartilhadas no sistema, como explicado em Zivot & Wang (2006). Os coeficientes ditam o grau de autorregressão das séries respectivamente. Os coeficientes são os coeficientes lineares do vetor de cointegração. As séries simuladas são geradas via um processo iterativo onde a cada nova iteração é gerada uma observação dos dados de baseado nos dados anteriores.

sendo: os coeficientes lineares dos autorregressivos definidos empiricamente como números aleatórios em um intervalo de [0.3, 0.9]; o coeficiente linear do termo de relação de longo prazo definidos empiricamente como números aleatórios em um intervalo de [0.01, 0.3]; uma variável com distribuição normal de média zero e desvio padrão igual a um.

Após gerar as séries cointegradas foram realizados dois testes de hipótese: um para estacionariedade e o outro para cointegração.

O teste de estacionariedade escolhido foi o Teste de Dickey–Fuller Aumentado (Hatanaka, 1996), mais conhecido como teste ADF. Para que o fluxo de modelagem continue é necessário que todas as séries simuladas sejam I(1), ou seja, integradas de ordem um. Caso essa condição seja satisfeita a um nível de confiança de 95% o algoritmo continua para a próxima etapa. Caso contrário é gerado um novo conjunto de dados cointegrados até que essa condição seja satisfeita.

A próxima etapa do processo é validar se as séries são cointegradas utilizando as estatísticas dos autovalores e autovetores de Johansen (Søren, 1991). Se a relação de cointegração não for satisfeita a nível de confiança de 95% é gerado um novo conjunto de dados cointegrados até que essa condição seja satisfeita. O intuito desta validação é selecionar apenas as séries que cumpririam as condições necessárias em uma análise empírica de cointegração.

# Validação dos Modelos

Para validar a capacidade preditiva dos modelos foi utilizada a metodologia de validação fora da amostra: foram separados conjuntos de dados de treino e teste para validar os resultados preditivos dos modelos variando o tamanho do período de teste conforme Figura 1.

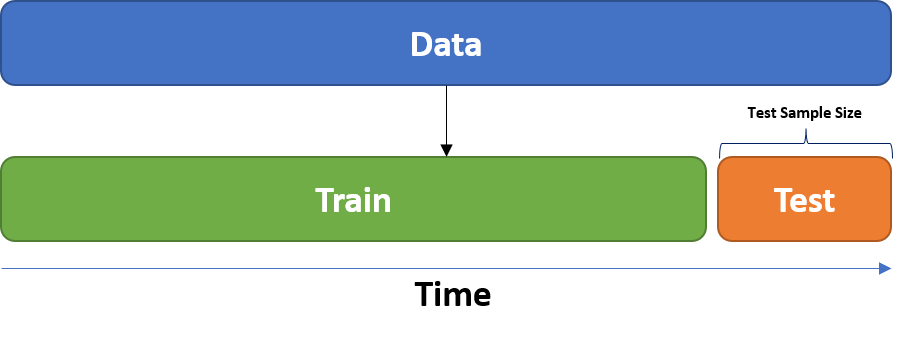


Figura - Estratégia de Validação Fora da Amostra Validation

A validação fora da amostra é essencial para saber se o modelo proposto está realmente generalizando de forma adequada aos dados ou está somente fazendo um ajuste nos dados de treinamento sem que este reflita em performance nos períodos de teste. O tamanho da amostra de teste será definido como um parâmetro da simulação. Para tal foram reservados entre 5%, 10% e 15% de todo o conjunto de dados mais recente para validar o modelo.

Após a construção das bases de treino e teste foram estimados os três modelos: ARIMA (1,1,0), VAR (1) em primeiras diferenças e um VEC (1). Gerados os objetos de modelos, foram realizadas as previsões e comparadas com os dados reservados para teste. Para mensurar a qualidade do ajuste foram utilizadas as seguintes métricas para avaliação da capacidade preditiva dos modelos:

* MAE (Mean Absolute Error);
* MSE (Mean Squared Error);
* Estatística de Diebold Mariano;
* Curvas de Distribuição Acumulada e distância de Kolmogorov-Smirnov (KS);

O MAE (Mean Absolute Error) é definido como erro absoluto médio, e pode ser calculado a partir da Equação 10.

Sendo: o número de observações preditas; a série de dados predita; a série de dados observada; os resíduos do modelo.

O RMSE (Root Mean Square Error) é definido como raiz do erro quadrático médio, e pode ser calculado a partir da Equação 11

As métricas acima (MAE, RMSE) não levam em consideração a direção dos erros somente a amplitude do termo de perturbação em relação à média. O MAE é menos sensível a valores atípicos. Por outro lado, o RMSE apesar de ser mais sensível a valores extremos pune com mais intensidade erros de maior magnitude.

Para gerar os intervalos de confiança das estatísticas estimadas foi utilizado o método de *Bootstrap* para gerar os intervalos de confiança empíricos descritos em (Mooney & Duval, 1993). Os parâmetros de *Bootstrap* utilizados foram: 1000 iterações; amostra de 20% da população por iteração e intervalos de confiança de 95%.

A Estatística de *Diebold-Mariano* (Diebold, 2013) compara previsões de dois modelos em períodos fora da amostra para checar se a acurácia entre ambos são estatisticamente iguais. O teste proposto inicialmente por (Diebold & Mariano, 1995) aplica uma função de custo sobre os erros de previsão de dois modelos e testa se o valor esperado , sendo definido como .

A curva de distribuição acumulada, ou ECDF, descreve como as probabilidades são associadas a valores de uma variável aleatória. A curva descreve a distribuição da probabilidade de uma variável aleatória e também é uma alternativa visual para observar as distribuições de dados entre as amostras de previsão. A partir da ECDF foi estimada a distância de KS (Daniel, Wayne W, 1990). A distância de Kolmogorov-Smirnov é usada como estatística no teste KS, que testa a probabilidade da hipótese nula de duas amostras serem retiradas da mesma distribuição. O objetivo é entender se a distribuição de erros entre os dois modelos testados é igual.

Após a validação das projeções os resultados são armazenados em bases de dados dedicadas. Esse processo se repete até que todas as combinações de parâmetros sejam testadas. O algoritmo completo de simulação desenvolvido para esse trabalho está descrito no fluxograma da Figura 2.

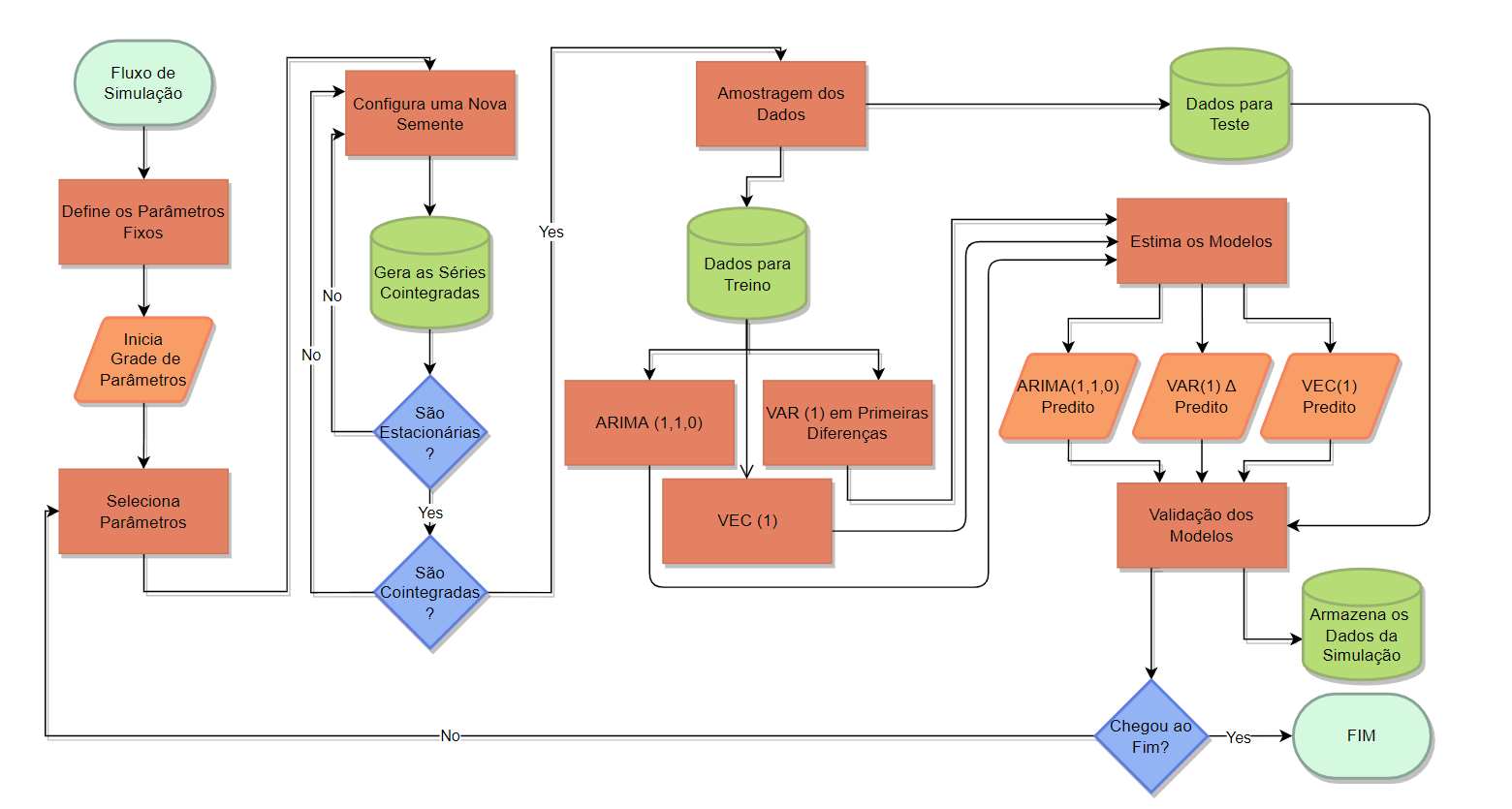


Figura - Fluxograma do Processo de Simulação

# Resultados

Ao fixar o número de observações em 100 unidades, sem uso de suavização por médias móveis e modelando o sistema com 3 variáveis foi observado que o modelo VEC apresenta níveis de MAE abaixo dos demais modelos. Quanto maior o período de treino do modelo menor é a média dos resíduos. A partir da Figura 3 observa-se que os resíduos reduzem na medida que o tamanho da amostra de teste também reduz de 85% para 90% e 95%.

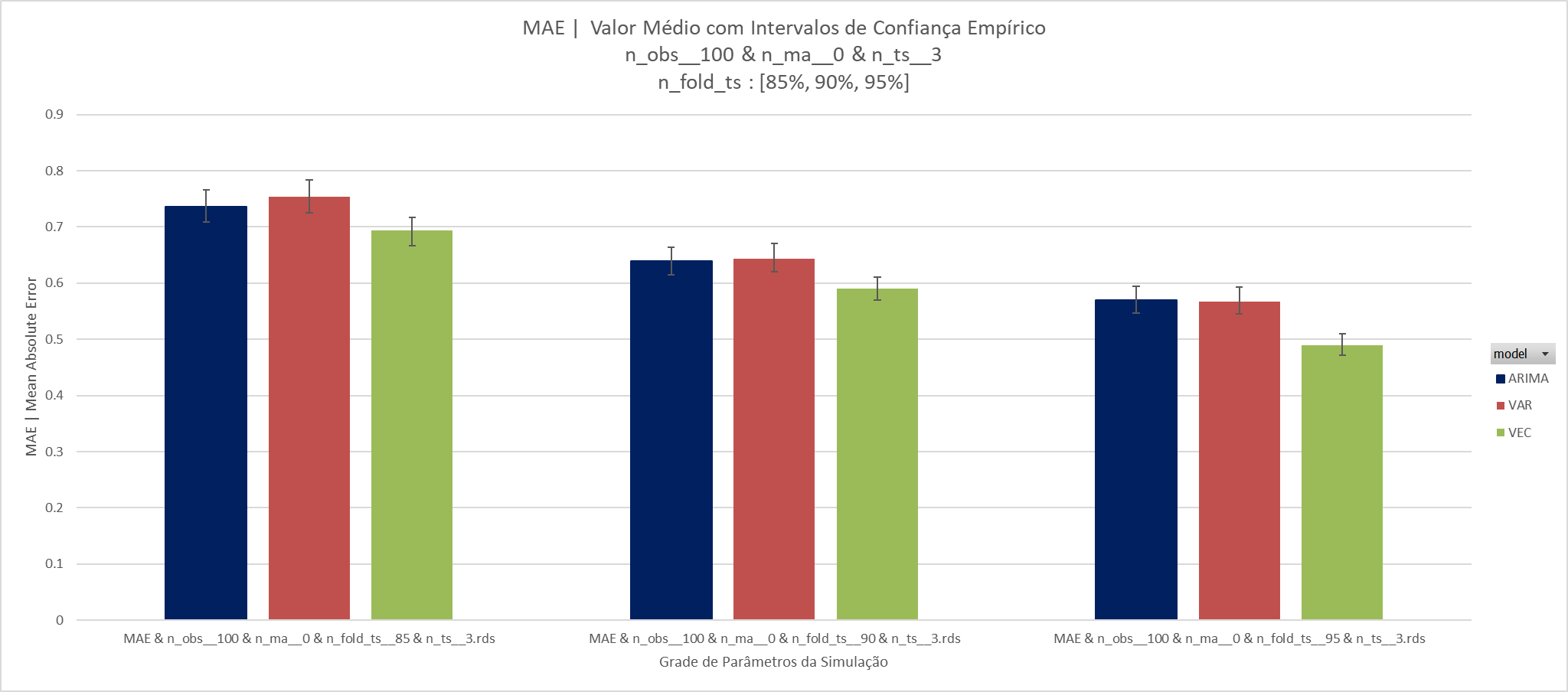


Figura – MAE | Fixo: n\_obs\_\_100 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90, 95]

O comportamento observado na Figura 3 ocorre também quando o número de observações da série aumenta de 100 para 250 e 500 unidades. Conforme Figura 4 e Figura 5:

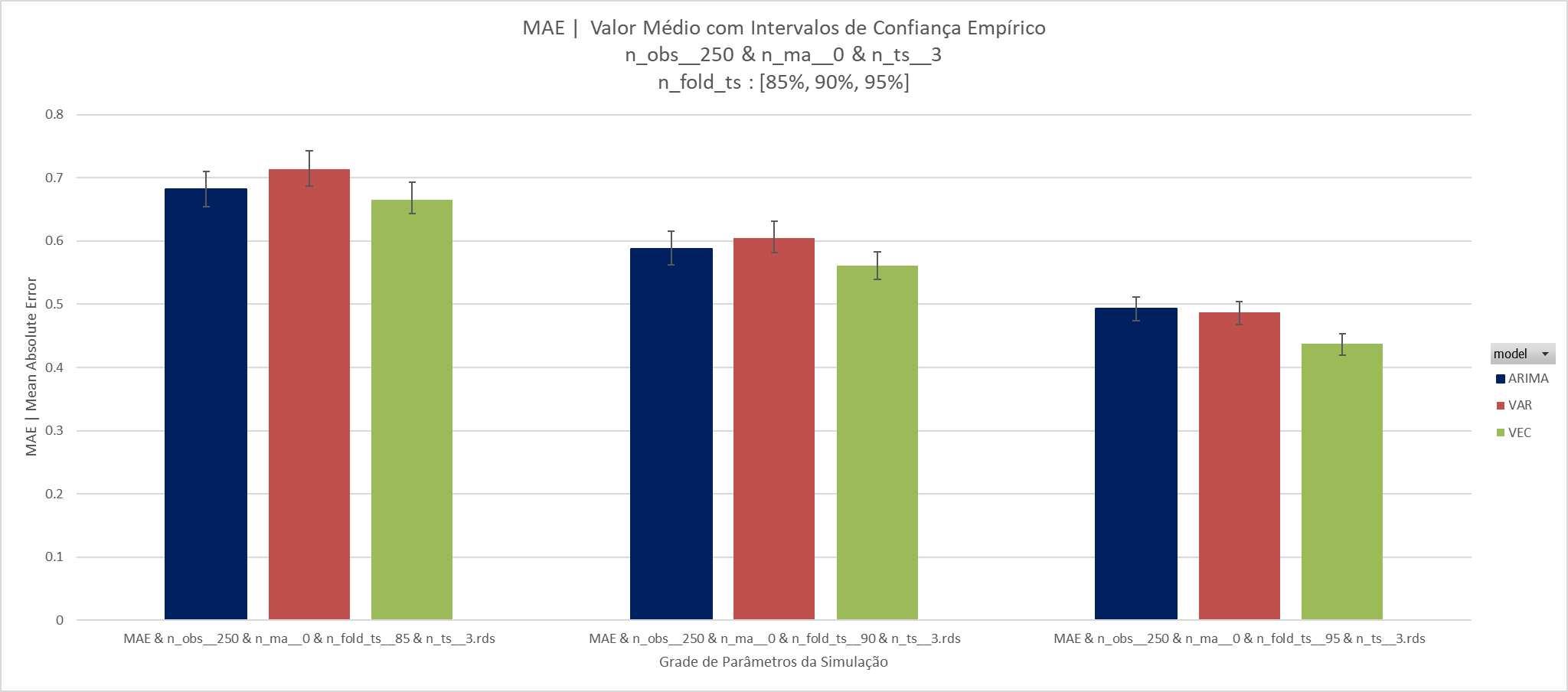


Figura - MAE | Fixo: n\_obs\_\_250 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90%, 95%]

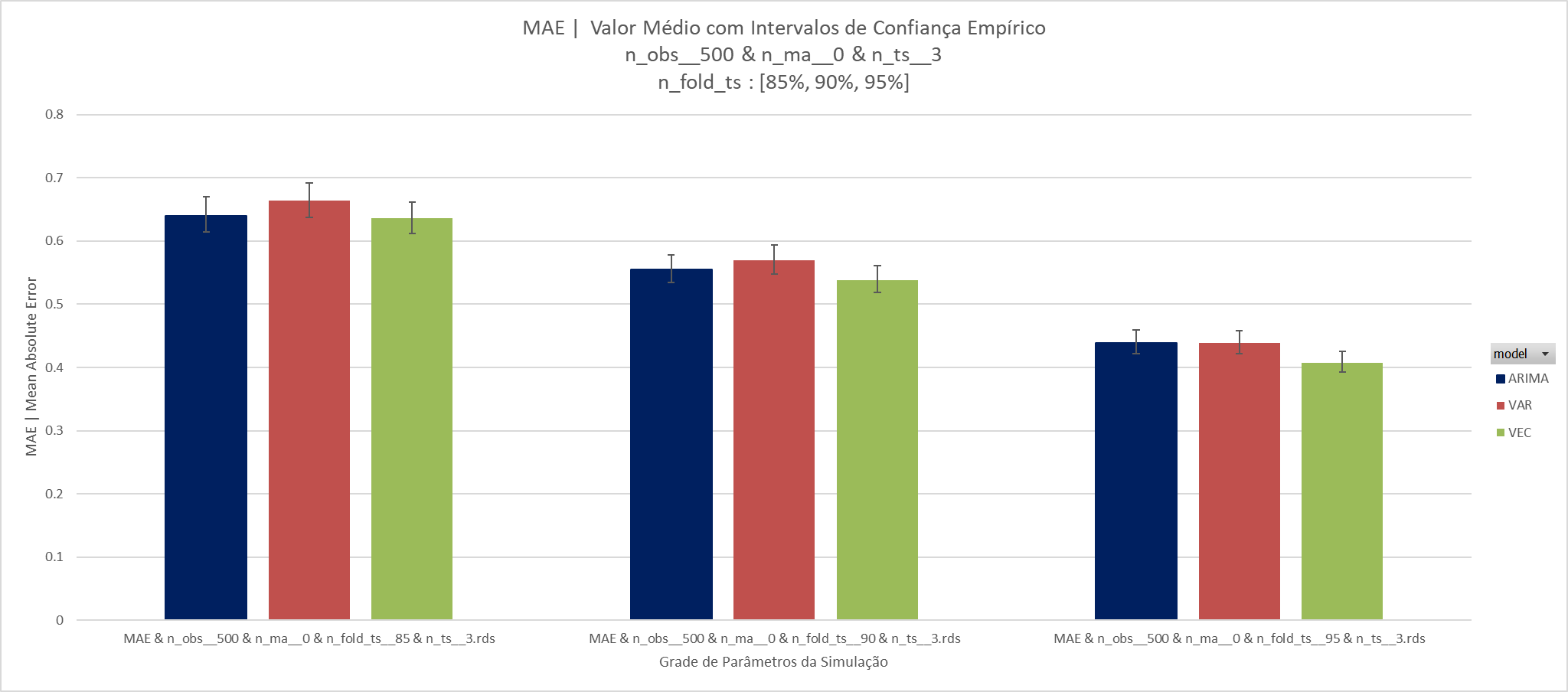


Figura - MAE | Fixo: n\_obs\_\_500 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90%, 95%]

Ao modelar séries de dados com 1000 observações observou-se que a performance dos modelos (treinados com até 85% da população) é igual dentro dos intervalos de confiança empíricos, conforme Figura 7. A principal hipótese formulada é que ao realizar previsões com horizontes de tempo acima de 150 observações a performance de todos os modelos é equivalente, ou seja, possivelmente todos convergiram a um valor médio e não é mais possível capturar diferenças de performance entre modelos.

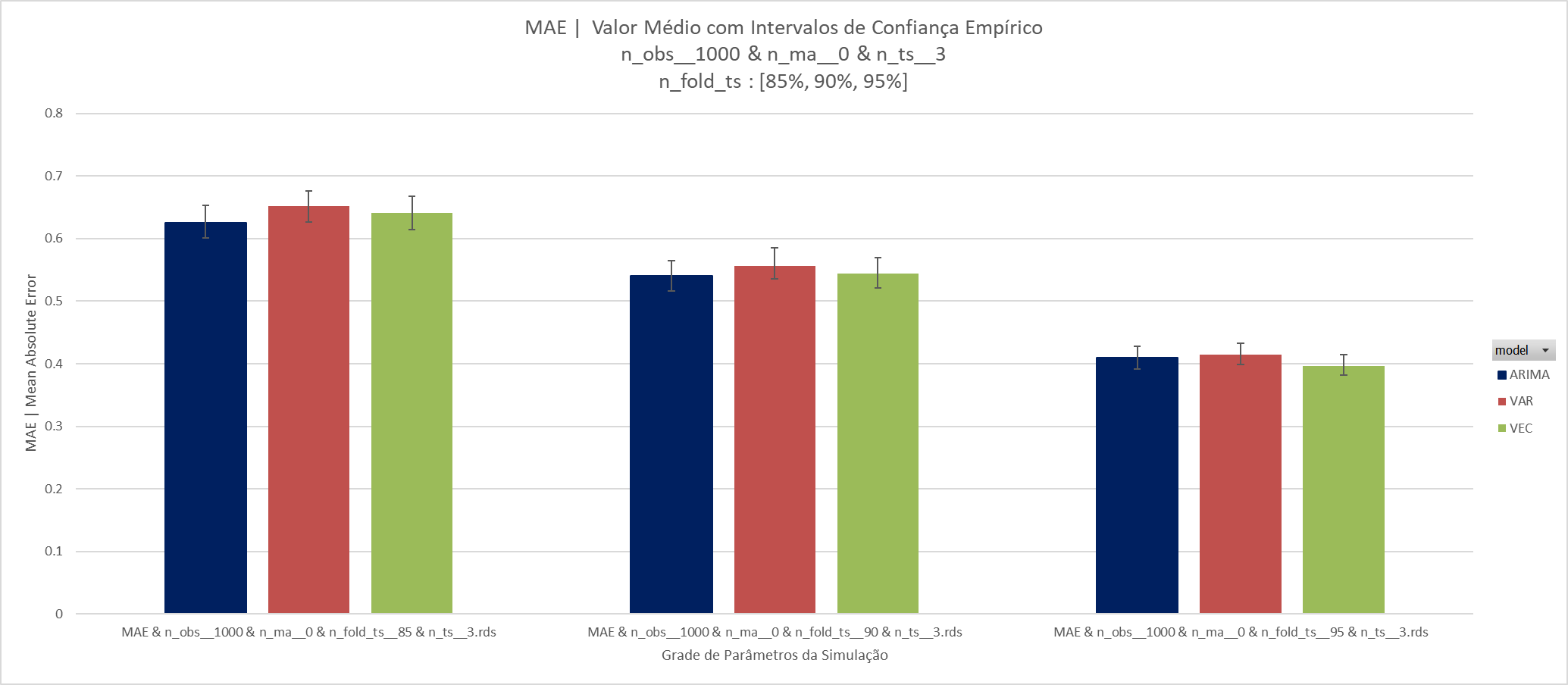


Figura - MAE | Fixo: n\_obs\_\_1000 & n\_ma\_\_0 & n\_ts\_\_3 | Variando: n\_fold\_ts: [85%, 90%, 95%]

Não foram encontradas diferenças significativas na performance dos modelos ao formular o problema com duas ou três variáveis. A Figura 7 mostra que os resultados obtidos nas simulações com 2 ou 3 variáveis é igual e contidas num mesmo intervalo de confiança. Esse comportamento ocorre tanto em séries temporais mais curtas (100 observações | Figura 7) quanto em séries mais longas (500 observações | Figura 8). Em ambos os casos modelo VEC obtém a melhor performance em dados fora da amostra.

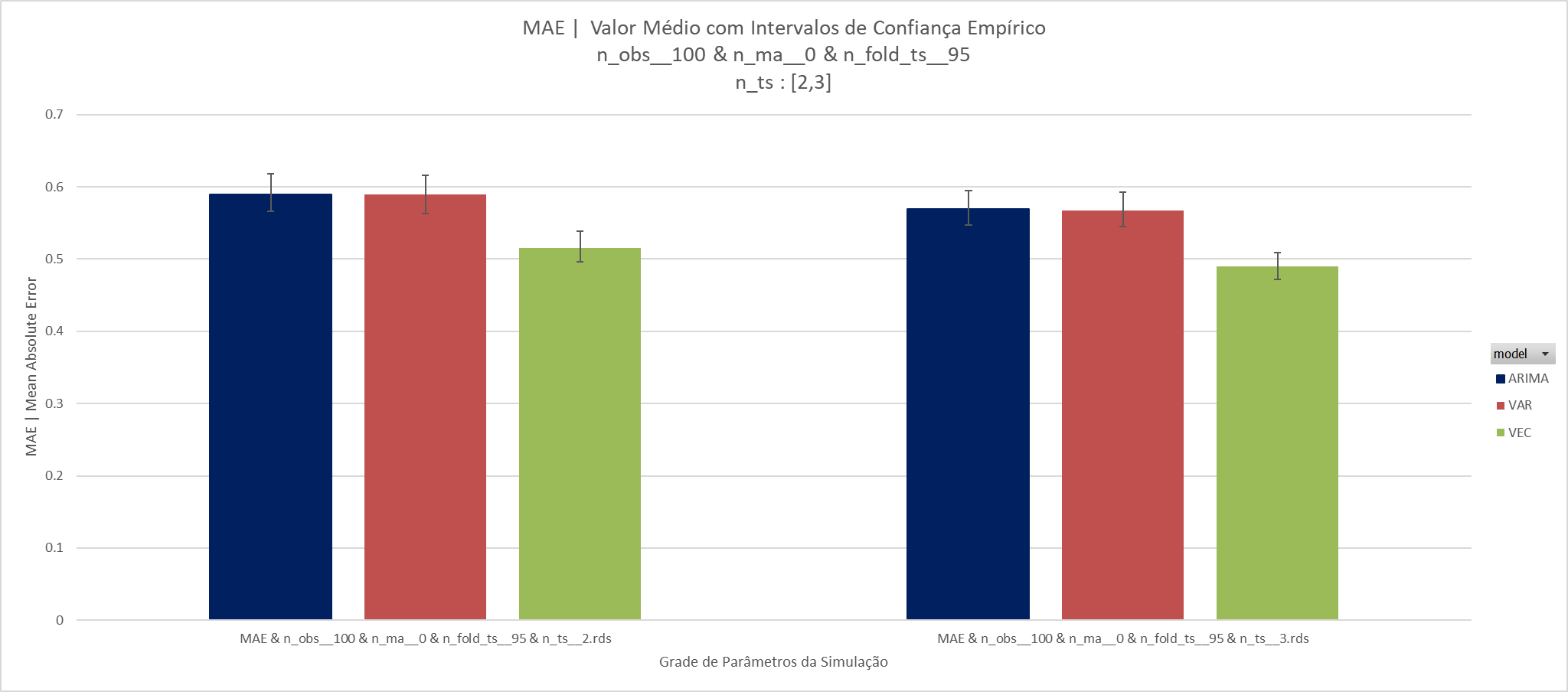


Figura - MAE | Fixo :n\_obs\_\_100 & n\_ma\_\_0 & n\_fold\_ts\_\_95 | Variando: n\_ts : [2,3]

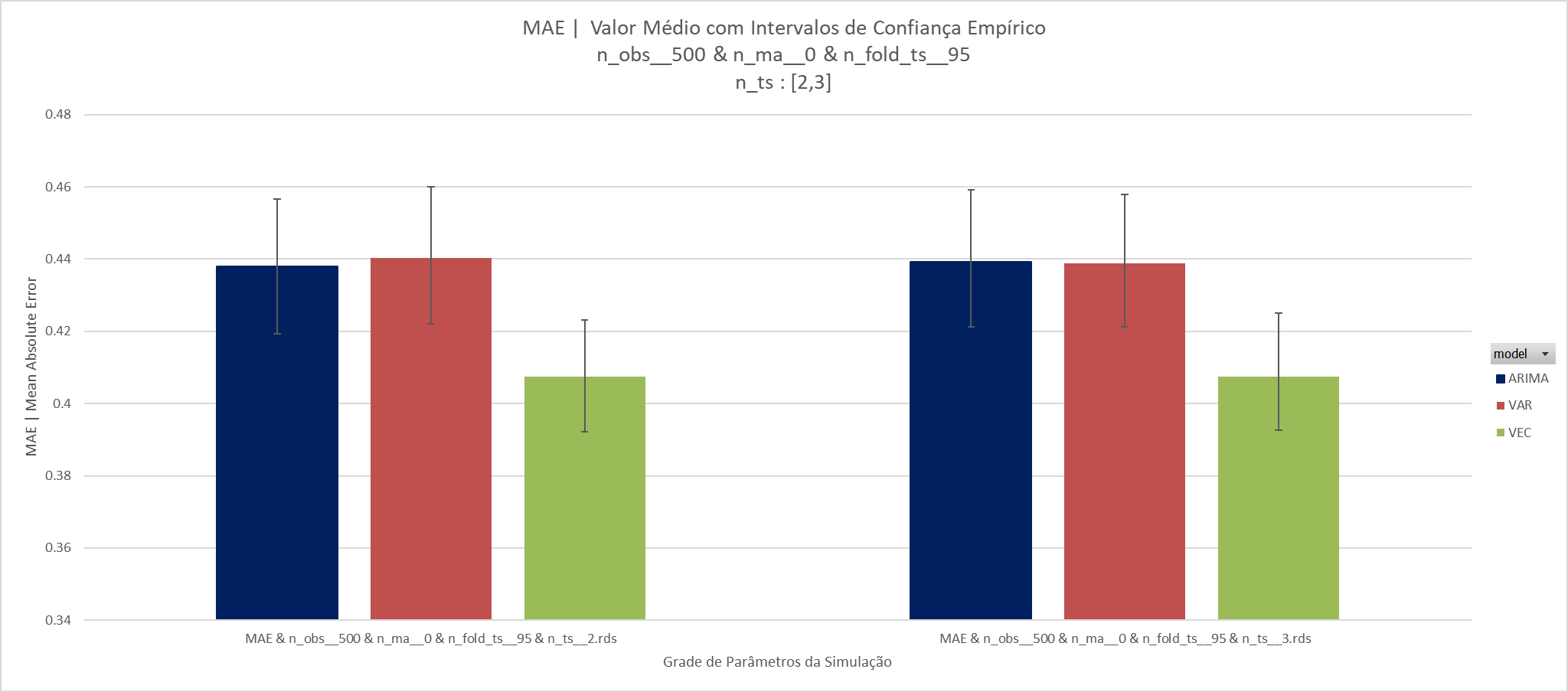


Figura - MAE | Fixo :n\_obs\_\_500 & n\_ma\_\_0 & n\_fold\_ts\_\_95 | Variando: n\_ts : [2,3]

Ao variar o número de observações em bases de dados fixando o período fora da amostra (sem uso de suavização por médias móveis e modelando o sistema com 3 variáveis) observa-se que quanto maior a amostra de teste menor é o nível de erros do modelo. Quanto mais dados o algoritmo recebe para treinamento menor é o patamar de erro dos modelos, algo que intuitivamente faz sentido. Entretanto, quanto maior o volume de observações utilizadas, menor é a distinção da capacidade preditiva entre modelos. Portanto quando se tem volume de dados grande para treinamento (acima de 500 observações) o modelo escolhido é menos relevante para as previsões de longo prazo. Na Figura 9 materializa as afirmações acima e observa-se que mesmo variando o tamanho da amostra o modelo VEC performa melhor que os demais.

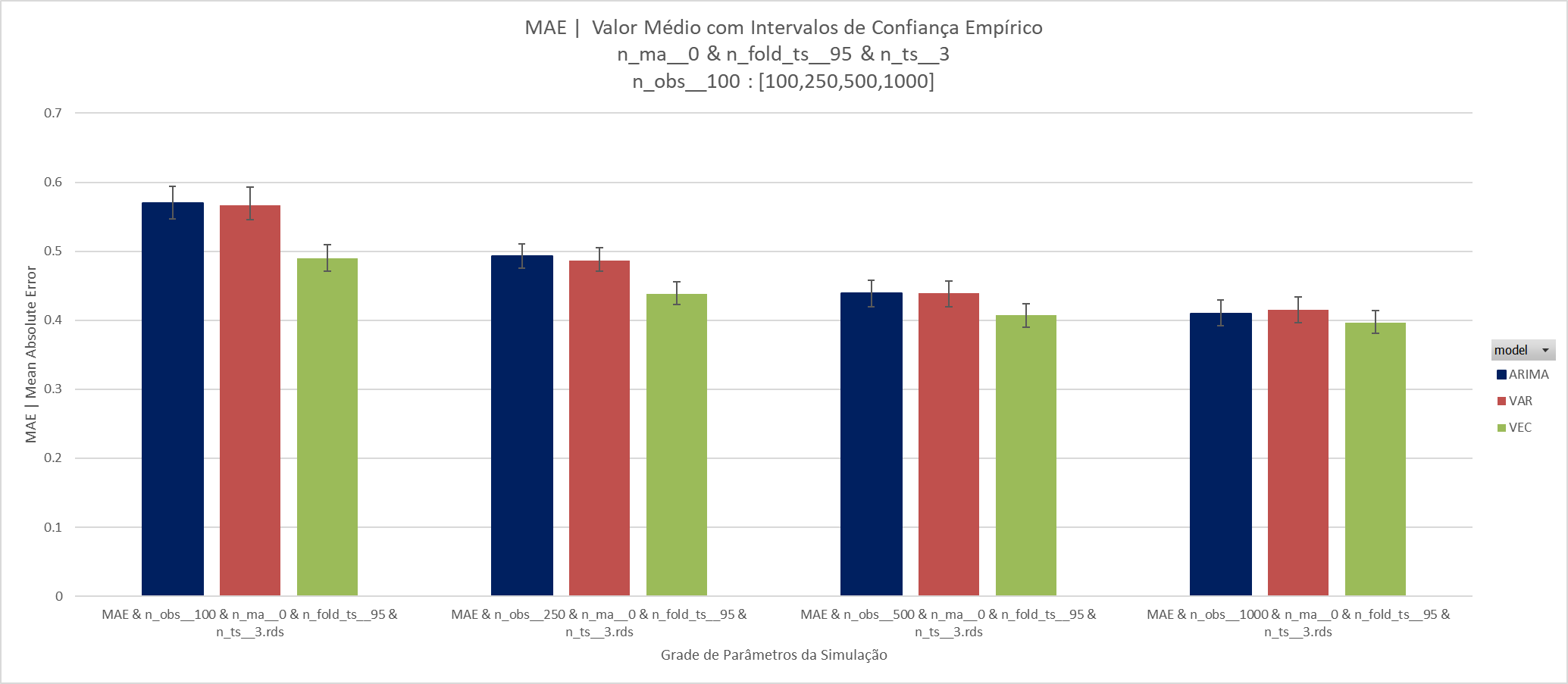


Figura - MAE | Fixo :n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_\_3 | Variando: n\_obs: [100,250,500,1000]

Ao aplicar uma suavização por média móvel de 12 períodos os modelos VEC apresentam performance igual ou inferior aos modelos ARIMA e VAR conforme Figura 10. O fato observado mostra que ao utilizar uma média móvel as séries são submetidas a uma espécie de filtro que remove componentes de alta frequência e que afetam diretamente a performance dos modelos vetoriais e principalmente o modelo VEC.

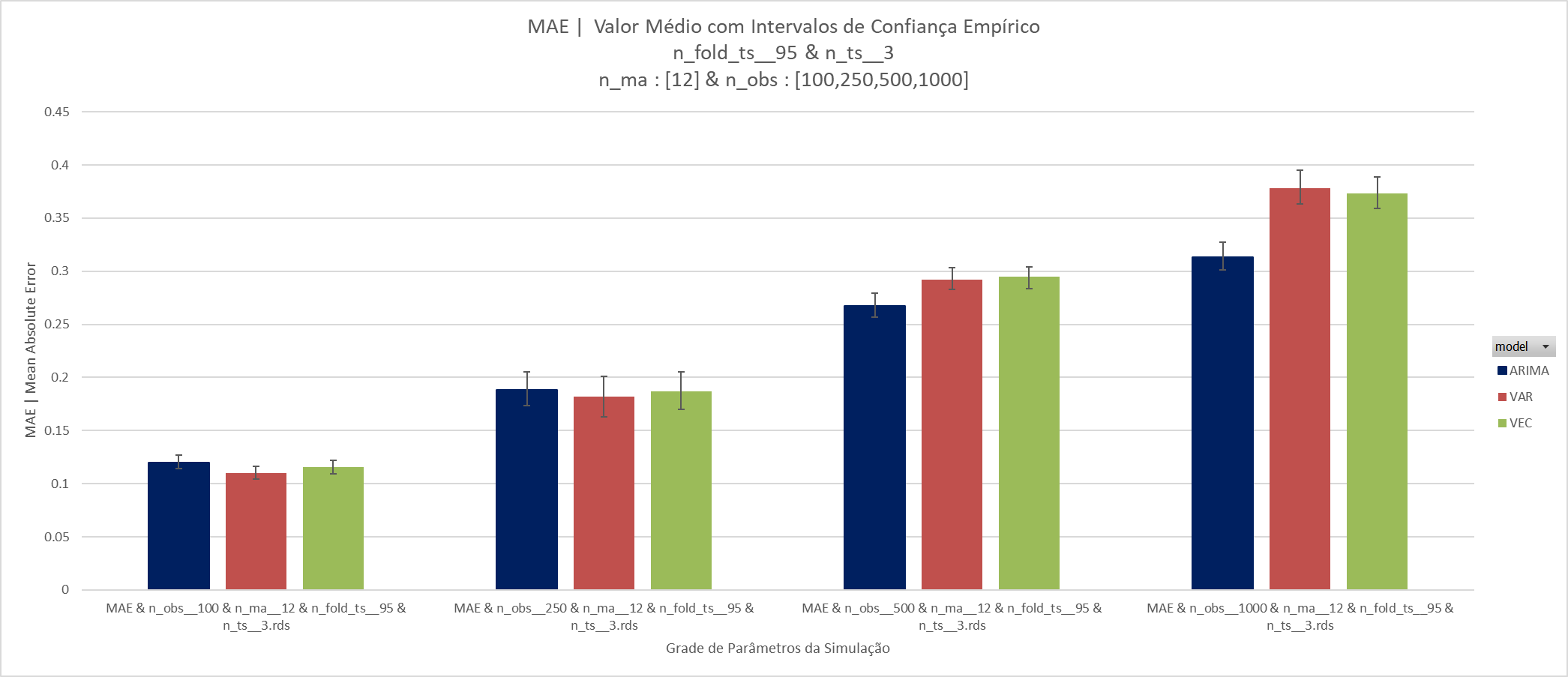


Figura - MAE | Fixo: n\_obs\_\_1000 & n\_fold\_ts\_\_95 & n\_ts\_\_3 | Variando: n\_ma: [0,12]

Os demais resultados encontrados na análise de erro médio absoluto (MAE) estão em Tabelas.

Os resultados observados ao validar os modelos com o erro médio quadrático (MSE) apresentam padrões de comportamento iguais aos observados para o MAE. Na Figura 11, assim como o observado para o MAE em Figura 9, observa-se que a magnitude dos erros cai à medida que são inseridos mais dados no modelo. Em ambas as análises (MAE e MSE) dentre os 3 modelos testados o que performa melhor é o VEC.

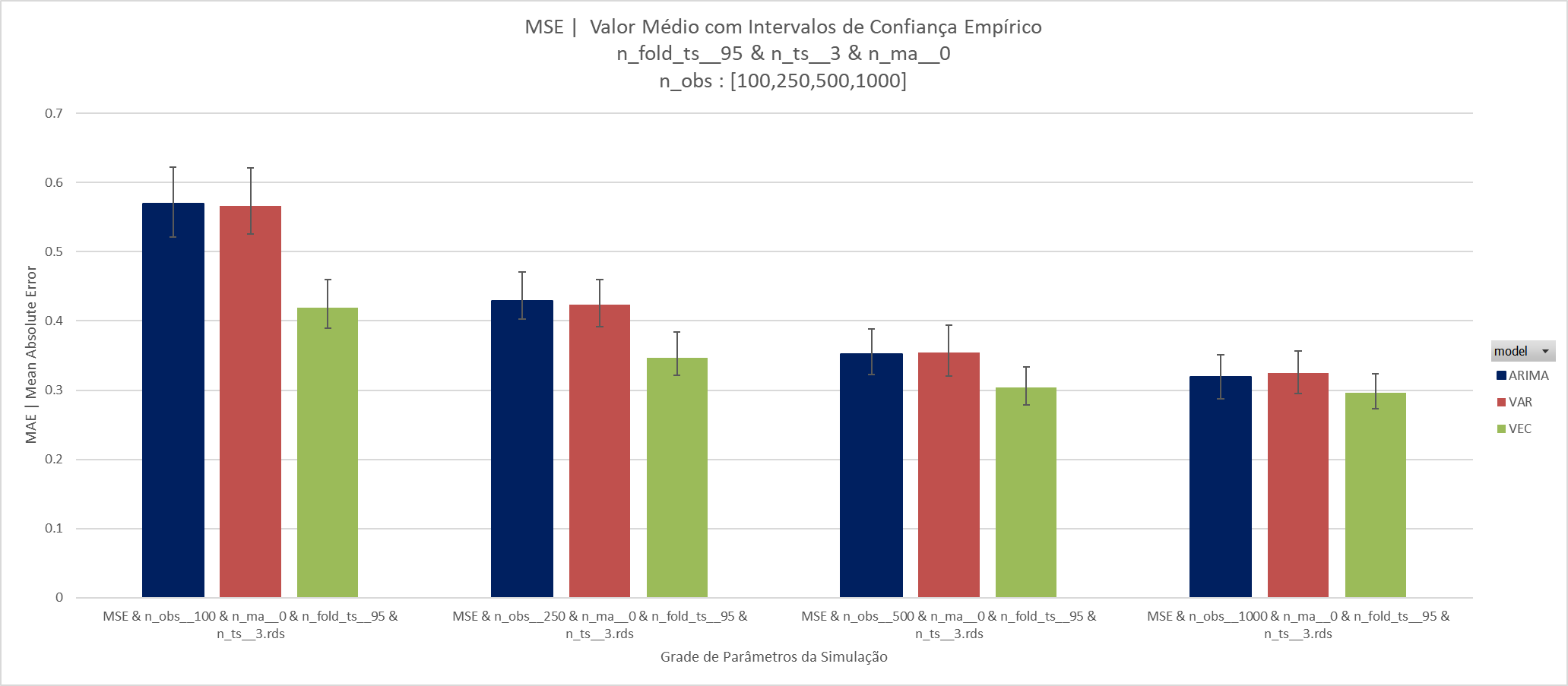


Figura - MSE | Fixo :n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_\_3 | Variando: n\_obs : [100,250,500,1000]

Com base nos resultados, percebeu-se que existe uma intersecção entre as interpretações obtidas na análise de MAE e MSE. Esses resultados não foram discutidos nesta sessão, mas são apresentados nas tabelas do anexo.

O teste de Diebold Mariano foi executado para todas as combinações de modelos e também utilizando 3 níveis de confiança: 90%, 95% e 99%. A hipótese nula testada é que a acurácia entre os modelos é estatisticamente igual em conjuntos de testes fora da amostra. Foi observado que quanto maior o período utilizado para teste (**n\_fold\_ts**) menor é a proporção de modelos que apresentam previsões iguais conforme Tabela 3.

Tabela - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_obs\_\_1000 & n\_ma\_\_0 & n\_fold\_ts\_\_85 & n\_ts\_3 | Variando: n\_fold\_ts: [85%,90%,95%]

|  |  |  |  |
| --- | --- | --- | --- |
|  | **% Modelos com acurácias iguais**  **Confiança: 99 %** | **% Modelos com acurácias iguais**  **Confiança: 95 %** | **% Modelos com acurácias iguais**  **Confiança: 90 %** |
| **n\_fold\_ts\n\_obs** | **1000** | **1000** | **1000** |
| **85%** |  |  |  |
| VEC vs ARIMA | 7% | 5% | 5% |
| VEC vs VAR | 7% | 6% | 5% |
| VAR vs ARIMA | 8% | 6% | 5% |
| **90%** |  |  |  |
| VEC vs ARIMA | 9% | 7% | 6% |
| VEC vs VAR | 9% | 6% | 5% |
| VAR vs ARIMA | 9% | 7% | 6% |
| **95%** |  |  |  |
| VEC vs ARIMA | 14% | 10% | 8% |
| VEC vs VAR | 14% | 11% | 8% |
| VAR vs ARIMA | 13% | 9% | 8% |

Foi observado também que quanto menor o número de observações da base de dados (**n\_obs**) menor é a distinção de acurácia dos modelos fora da amostra conforme Tabela 4.

Tabela - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_obs\_\_250 & n\_ma\_\_0 & n\_fold\_ts\_\_85 & n\_ts\_3 | Variando: n\_fold\_ts: [85%,90%,95%]

|  |  |  |  |
| --- | --- | --- | --- |
|  | **% Modelos com acurácias iguais**  **Confiança: 99 %** | **% Modelos com acurácias iguais**  **Confiança: 95 %** | **% Modelos com acurácias iguais**  **Confiança: 90 %** |
| **n\_fold\_ts\n\_obs** | **250** | **250** | **250** |
| **85%** |  |  |  |
| VEC vs ARIMA | 10% | 7% | 7% |
| VEC vs VAR | 13% | 10% | 8% |
| VAR vs ARIMA | 10% | 7% | 7% |
| **90%** |  |  |  |
| VEC vs ARIMA | 13% | 9% | 8% |
| VEC vs VAR | 13% | 10% | 10% |
| VAR vs ARIMA | 12% | 8% | 7% |
| **95%** |  |  |  |
| VEC vs ARIMA | 17% | 11% | 8% |
| VEC vs VAR | 21% | 14% | 11% |
| VAR vs ARIMA | 19% | 13% | 9% |

Ao variar o tamanho das séries de dados (1000, 500 e 250 observações) mantendo fixo o percentual fora da amostra em 95% dos dados tem-se que quanto menor a amostra total (treino + teste) maior é a proporção de modelos com acurácia igual em períodos fora da amostra.

Tabela - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_3 | Variando: n\_obs: [100,500,250]

|  |  |  |  |
| --- | --- | --- | --- |
|  | **% Modelos com acurácias iguais**  **Confiança: 99 %** | **% Modelos com acurácias iguais**  **Confiança: 95 %** | **% Modelos com acurácias iguais**  **Confiança: 90 %** |
| **n\_obs\n\_fold\_ts** | **95%** | **95%** | **95%** |
| **1000** |  |  |  |
| VEC vs ARIMA | 14% | 10% | 8% |
| VEC vs VAR | 14% | 11% | 8% |
| VAR vs ARIMA | 13% | 9% | 8% |
| **500** |  |  |  |
| VEC vs ARIMA | 17% | 11% | 8% |
| VEC vs VAR | 21% | 14% | 11% |
| VAR vs ARIMA | 19% | 13% | 9% |
| **250** |  |  |  |
| VEC vs ARIMA | 55% | 21% | 14% |
| VEC vs VAR | 57% | 27% | 18% |
| VAR vs ARIMA | 53% | 23% | 14% |

De forma já observada anteriormente nas análises de MAE, ao variar o número de séries temporais no sistema não foi observado diferenças relevantes de acurácia dos modelos em conjuntos fora da amostra ao comparar as simulações usando duas ou três variáveis no sistema de equações, conforme Tabela 6

Tabela - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_obs\_\_500 | Variando: n\_ts: [2,3]

|  |  |  |  |
| --- | --- | --- | --- |
|  | **% Modelos com acurácias iguais**  **Confiança: 99 %** | **% Modelos com acurácias iguais**  **Confiança: 95 %** | **% Modelos com acurácias iguais**  **Confiança: 90 %** |
| **n\_ts\n\_fold\_ts** | **95%** | **95%** | **95%** |
| **2** |  |  |  |
| VEC vs ARIMA | 21% | 13% | 10% |
| VEC vs VAR | 22% | 15% | 12% |
| VAR vs ARIMA | 19% | 13% | 10% |
| **3** |  |  |  |
| VEC vs ARIMA | 17% | 11% | 8% |
| VEC vs VAR | 21% | 14% | 11% |
| VAR vs ARIMA | 19% | 13% | 9% |

Ao utilizar o tratamento de suavização por médias móveis é observado que o percentual de modelos que apresentam acurácia igual em conjunto de dados fora da amostra de treino é menor do que a série original conforme Tabela 7.

Tabela - Percentual de Modelos com Performance igual fora da amostra segundo teste de Diebold Mariano | Fixando: n\_ts\_\_3 & n\_fold\_ts\_\_95 & n\_obs\_\_500 | Variando: n\_ma: [0,12]

|  |  |  |  |
| --- | --- | --- | --- |
|  | **% Modelos com acurácias iguais**  **Confiança: 99 %** | **% Modelos com acurácias iguais**  **Confiança: 95 %** | **% Modelos com acurácias iguais**  **Confiança: 90 %** |
| **n\_ma\n\_fold\_ts** | **95%** | **95%** | **95%** |
| **0** |  |  |  |
| VEC vs ARIMA | 39% | 29% | 23% |
| VEC vs VAR | 34% | 26% | 20% |
| VAR vs ARIMA | 36% | 27% | 23% |
| **12** |  |  |  |
| VEC vs ARIMA | 17% | 11% | 8% |
| VEC vs VAR | 21% | 14% | 11% |
| VAR vs ARIMA | 19% | 13% | 9% |

Ao avaliar os resultados do teste KS sobre os erros MAE observa-se que para quase todas as combinações de parâmetros sem médias móveis não foi possível rejeitar (com 95% de confiança) a hipótese nula de que os resultados de performance dos modelos foram retirados de uma mesma amostra. A partir da Tabela 8 pode-se afirmar que todos os modelos apresentam comportamentos de performance fora da amostra igual com as premissas do teste de hipótese de KS.

Tabela – Teste de KS para cada combinação de parâmetros da simulação mantendo fixo o número de observações em 500 unidades.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parâmetros da Simulação** | | | | **Nível de Confiança: 95%**  **MAE : Retirados da Mesma Amostra?** | | |
| **n\_obs** | **n\_ma** | **n\_fold\_ts** | **n\_ts** | **ARIMA vs VAR** | **ARIMA vs VEC** | **VAR vs VEC** |
| 500 | 0 | 85 | 2 | VERDADEIRO | VERDADEIRO | VERDADEIRO |
| 500 | 0 | 85 | 3 | VERDADEIRO | VERDADEIRO | VERDADEIRO |
| 500 | 0 | 90 | 2 | VERDADEIRO | VERDADEIRO | FALSO |
| 500 | 0 | 90 | 3 | VERDADEIRO | VERDADEIRO | VERDADEIRO |
| 500 | 0 | 95 | 2 | VERDADEIRO | VERDADEIRO | VERDADEIRO |
| 500 | 0 | 95 | 3 | VERDADEIRO | VERDADEIRO | VERDADEIRO |
| 500 | 12 | 95 | 3 | VERDADEIRO | FALSO | VERDADEIRO |

Os gráficos de ECDF da Figura 12 corroboram com o resultado de teste apresentado na Tabela 8.

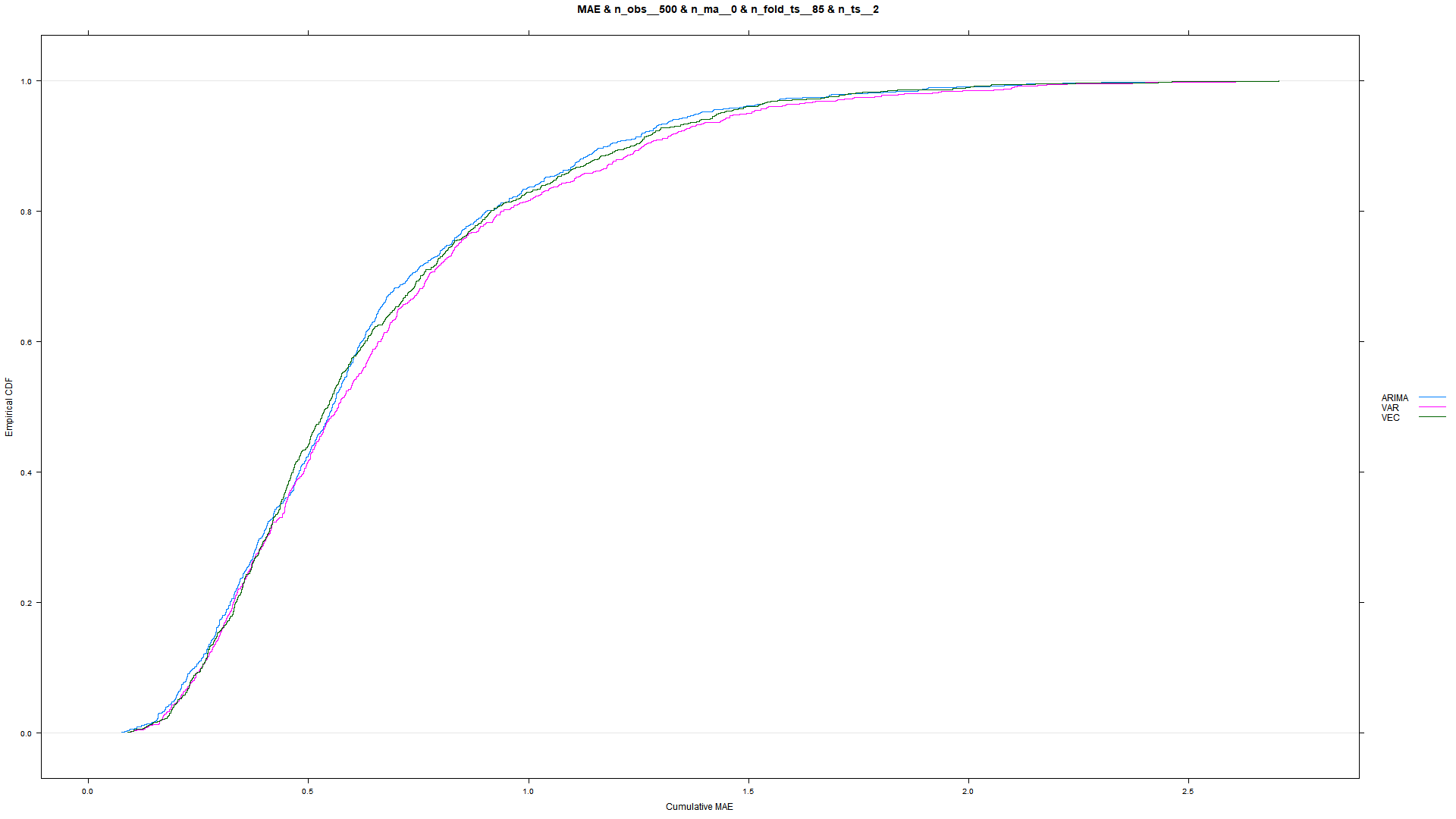


Figura – ECDF | MAE & n\_obs\_\_500 & n\_ma\_\_0 & n\_fold\_ts\_\_85 & n\_ts\_\_2

Quando o tamanho da amostra utilizada para treino do modelo é menor (100 unidades) tem-se que os modelos ARIMA e VAR apresentam uma performance semelhante e o modelo VEC se destaca apresentando performance fora da amostra superior aos demais conforme Tabela 9 e Figura 13.

Tabela – Teste de KS para cada combinação de parâmetros da simulação mantendo fixo o número de observações em 100 unidades.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parâmetros da Simulação** | | | | **Nível de Confiança: 95%**  **MAE : Retirados da Mesma Amostra?** | | |
| **n\_obs** | **n\_ma** | **n\_fold\_ts** | **n\_ts** | **ARIMA vs VAR** | **ARIMA vs VEC** | **VAR vs VEC** |
| 100 | 0 | 85 | 2 | VERDADEIRO | VERDADEIRO | FALSO |
| 100 | 0 | 85 | 3 | VERDADEIRO | VERDADEIRO | VERDADEIRO |
| 100 | 0 | 90 | 2 | VERDADEIRO | FALSO | FALSO |
| 100 | 0 | 90 | 3 | VERDADEIRO | FALSO | FALSO |
| 100 | 0 | 95 | 2 | VERDADEIRO | FALSO | FALSO |
| 100 | 0 | 95 | 3 | VERDADEIRO | FALSO | FALSO |

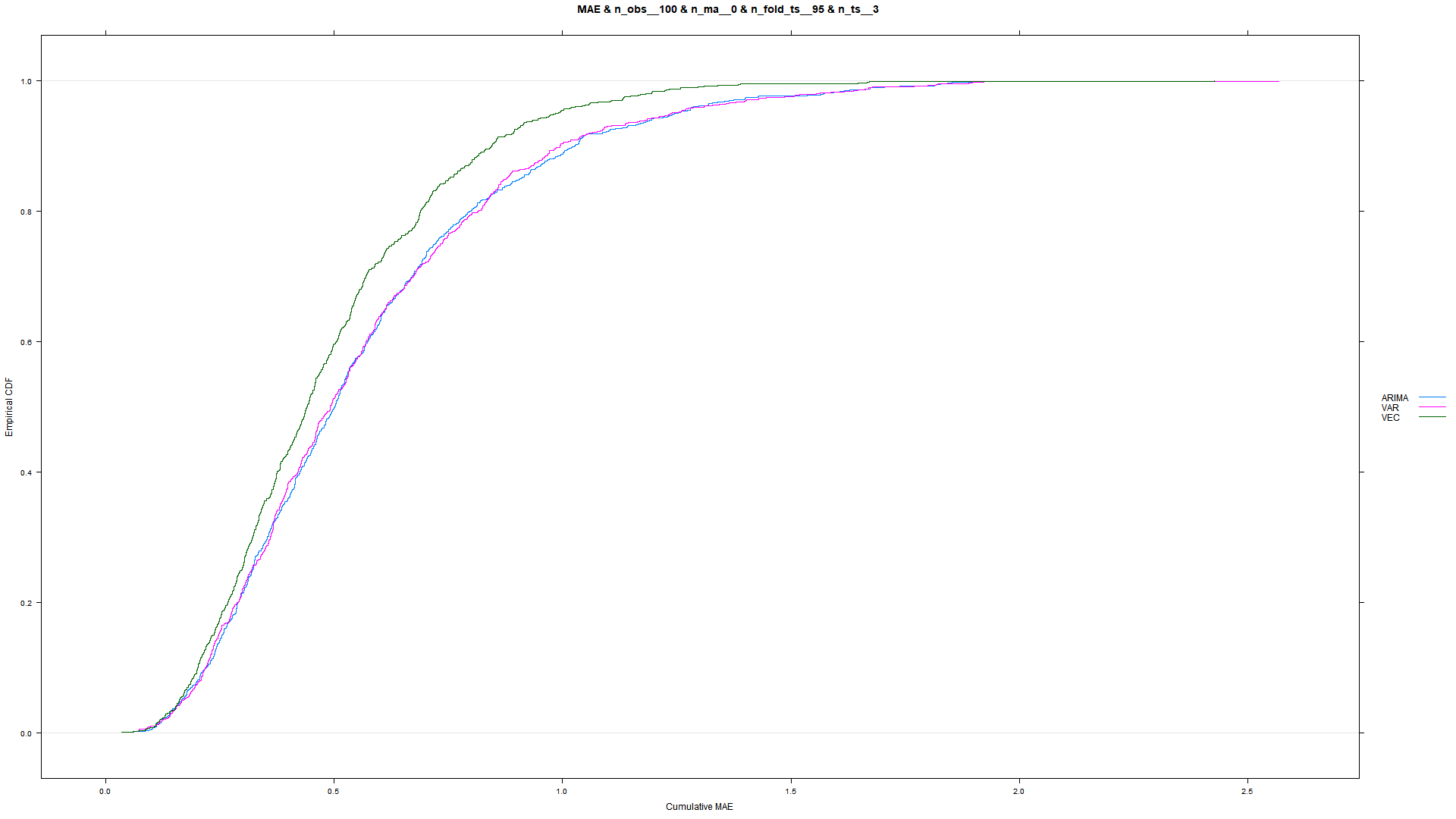


Figura - ECDF | MAE & n\_obs\_\_100 & n\_ma\_\_0 & n\_fold\_ts\_\_95 & n\_ts\_\_3

# Conclusões

O modelo VEC apresentou níveis de erro abaixo dos modelos VAR e ARIMA ao modelar séries de dados cointegradas. O comportamento observado ocorreu tanto em amostras de 100 até 1000 observações.

Quanto maior o período de treinamento do modelo menor é a média dos resíduos. Modelos treinados com séries de dados longas (acima de 500 observações) não apresentaram resíduos estatisticamente distintos, os níveis de erro encontrados parecem convergir para uma mesma distribuição e média.

O número de séries temporais no sistema não afetou a performance individual dos modelos. Conforme esperado, observou-se que quanto maior a amostra de teste menor é o nível de erros do modelo.

Ao aplicar uma suavização por médias móveis de 12 períodos, os modelos VEC apresentaram performance igual ou inferior aos modelos ARIMA e VAR, fato esse, explicado pela remoção de componentes de mais alta frequência que são modelados com maior precisão em modelos VEC.

A validação de modelos com MAE e MSE são equivalentes; ambos convergiram para resultados de mesma interpretação.

Quanto menor o número de observações da base de dados menor é a distinção de acurácia dos modelos fora da amostra de treinamento segundo o teste de Diebold Mariano.

Contudo, os intervalos de confiança das estatísticas descritivas alertam para uma necessidade maior de testes e de validações mais complexas para que os analistas/cientistas optem por uma estrutura em detrimento de outra. A simulação e toda validação apresentada reforçam o fato de que é sempre necessário o uso de um conjunto grande de técnicas de validação para a seleção de modelos.

Por fim, conclui-se que no conjunto particular de séries cointegradas estudadas neste trabalho, os modelos VEC apresentaram performance preditiva superior em dados fora da amostra. No entanto, seus intervalos de confiança refletem para uma possível equivalência entre os modelos em horizontes longos de previsão.

# Referências

Box, G.E.P. and Jenkins, G.M. (1970) Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Lütkepohl, H. (1991). Introduction to Multiple Time Series Analysis. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-02691-5>

Morettin, P. A. and Toloi. C. M. (2004). Análise de Séries Temporais. Edgard Blücher, São Paulo.

Enders, W. (2014) Applied Econometric Time Series. 4th Edition. John Wiley, New York.

Daniel, Wayne W. (1990). "Kolmogorov–Smirnov one-sample test". Applied Nonparametric Statistics (2nd ed.). Boston: PWS-Kent. pp. 319–330. ISBN 978-0-534-91976-4.

Hatanaka, Michio (1996). Time-Series-Based Econometrics: Unit Roots and Cointegration. New York: Oxford University Press. pp. 48–49. ISBN 978-0-19-877353-5.

Mooney, C.Z. & Duval, R.D. (1993). Bootstrapping: A nonparametric approach to statistical inference. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-095. Newbury Park, CA: Sage.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Avaliable in: <https://www.R-project.org> (Accessed on Mar 10, 2022).

Zivot, E. & Wang, J. (2006) “Modeling Financial Time Series With S-Plus”. Springer Science + Business Media, Inc.

Abdullah, L., & Dwivedy, N. (2012). ARIMA Model for Gold Bullion Coin Selling Prices Forecasting Cite this paper Related papers GOLD PRICE FORECAST ING IN INDIA USING ARIMA MODELLING AARF Publicat ions Journals, SAINA BABY Capt uring Volat ilit y Wit h St at a t o forecast Coriander Price b. *International Journal of Advances in Applied Sciences (IJAAS)*, *1*(4), 153–158.

Afana, M., Ahmed, J., Harb, B., Abu-Nasser, B. S., & Abu-Naser, S. S. (2018). Artificial Neural Network for Forecasting Car Mileage per Gallon in the City. *International Journal of Advanced Science and Technology*, *124*, 51–59. http://dx.doi.org/10.14257/ijast.2018.124.05

Anggraeni, W., Boga, K., & Mahananto, F. (2018). ScienceDirect ScienceDirect The Performance of ARIMAX Model and Vector Autoregressive ( VAR ) Model in Forecasting Strategic Commodity Price in Indonesia. *Procedia Computer Science*, *124*, 189–196. https://doi.org/10.1016/j.procs.2017.12.146

Araz, O. M., Choi, T. M., Olson, D. L., & Salman, F. S. (2020). Data Analytics for Operational Risk Management. *Decision Sciences*, *51*(6), 1316–1319. https://doi.org/10.1111/deci.12443

Bachmeier, L. J., & Swanson, N. R. (2005). Predicting inflation: Does the quantity theory help? *Economic Inquiry*, *43*(3), 570–585. https://doi.org/10.1093/ei/cbi039

Chen, T., Xie, L., & Zhang, Y. (2017). How does analysts’ forecast quality relate to corporate investment efficiency? *Journal of Corporate Finance*, *43*, 217–240. https://doi.org/10.1016/j.jcorpfin.2016.12.010

Diebold, F. X. (2013). *Comparing Predictive Accuracy , Twenty Years Later : A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests ∗ Comparing Model-Free Forecasts*.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Diebold, F. X., Rudebusch, G. D., & Boraǧan Aruoba, S. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, *131*(1–2), 309–338. https://doi.org/10.1016/j.jeconom.2005.01.011

Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*(2), 251. https://doi.org/10.2307/1913236

Fanchon, P., & Wendel, J. (1992). Estimating VAR models under non-stationarity and cointegration: Alternative approaches for forecasting cattle prices. *Applied Economics*, *24*(2), 207–217. https://doi.org/10.1080/00036849200000119

Hargreaves, J. C., Annan, J. D., Edwards, N. R., & Marsh, R. (2004). An efficient climate forecasting method using an intermediate complexity Earth System Model and the ensemble Kalman filter. *Climate Dynamics*, *23*(7–8), 745–760. https://doi.org/10.1007/s00382-004-0471-4

Hu, Z., Ge, Q., Li, S., & Xiong, M. (2020). Artificial Intelligence Forecasting of Covid-19 in China. *International Journal of Educational Excellence*, *6*(1), 71–94. https://doi.org/10.18562/ijee.054

Ioannidis, J. P. A., Cripps, S., & Tanner, M. A. (2020). Forecasting for COVID-19 has failed. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2020.08.004

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*(2–3), 231–254. https://doi.org/10.1016/0165-1889(88)90041-3

Jones, S. A., Joy, M. P., & Pearson, J. (2002). Forecasting demand of emergency care. *Health Care Management Science*, *5*(4), 297–305. https://doi.org/10.1023/A:1020390425029

Mohamed, J. (2020). *Time Series Modeling and Forecasting of Somaliland Consumer Price Index : A Comparison of ARIMA and Regression with ARIMA Errors*. *9*(4), 143–153. https://doi.org/10.11648/j.ajtas.20200904.18

Ng, S. T., Fan, R. Y. C., & Wong, J. M. W. (2011). An econometric model for forecasting private construction investment in Hong Kong. *Construction Management and Economics*, *29*(5), 519–534. https://doi.org/10.1080/01446193.2011.570356

Nyoni, T. (2018). Modeling and Forecasting Naira / USD Exchange Rate In Nigeria: a Box - Jenkins ARIMA approach. *Mrpa*, *MPRA Paper No. 88622*, 6–25. https://mpra.ub.uni-muenchen.de/88622/

Rogalski, R. J., & Vinso, J. D. (1977). Stock Returns , Money Supply and the Direction of Causality. *American Finance Association*, *32*(4), 1017–1030.

Salisu, A. A., Ebuh, G. U., & Usman, N. (2020). Revisiting oil-stock nexus during COVID-19 pandemic: Some preliminary results. *International Review of Economics and Finance*, *69*(June), 280–294. https://doi.org/10.1016/j.iref.2020.06.023

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*(1), 1. https://doi.org/10.2307/1912017

Søren, J. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, *59*(6), 1551–1580. https://doi.org/10.1017/CBO9781107415324.004

Tehreem, H. S., Anser, M. K., Nassani, A. A., Abro, M. M. Q., & Zaman, K. (2020). Impact of average temperature, energy demand, sectoral value added, and population growth on water resource quality and mortality rate: it is time to stop waiting around. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-020-09822-w

Wangdi, K., Singhasivanon, P., Silawan, T., Lawpoolsri, S., White, N. J., & Kaewkungwal, J. (2010). Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: A case study in endemic districts of Bhutan. *Malaria Journal*, *9*(1), 1–9. https://doi.org/10.1186/1475-2875-9-251

# Apêndice

# Repositório

Todos os scripts utilizados para confecção deste trabalho e resultados consolidados estão armazenados no seguinte repositório do GitHub: <https://github.com/ThiagueraBarao/master_degree>.